

University of Groningen

Applying organizational justice theory to admission into higher education

Niessen, A. Susan M.; Meijer, Rob R.; Tendeiro, Jorge N.

Published in:
International Journal of Selection and Assessment

DOI:
[10.1111/ijsa.12161](https://doi.org/10.1111/ijsa.12161)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017). Applying organizational justice theory to admission into higher education: Admission from a student perspective. *International Journal of Selection and Assessment*, 25(1), 72-84. <https://doi.org/10.1111/ijsa.12161>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Applying organizational justice theory to admission into higher education: Admission from a student perspective

A. Susan M. Niessen | Rob R. Meijer | Jorge N. Tendeiro

Department of Psychometrics and Statistics,
Faculty of Behavioral and Social Sciences,
University of Groningen, Grote Kruisstraat
2/1, 9712 TS Groningen, The Netherlands

Correspondence

A. Susan M. Niessen, Department of
Psychometrics and Statistics, Faculty of
Behavioral and Social Sciences, University
of Groningen, Grote Kruisstraat 2/1, 9712
TS Groningen, The Netherlands.
Email: a.s.m.niessen@rug.nl

Abstract

Applicant perceptions of methods used in admission procedures to higher education were investigated using organizational justice theory. Applicants to a psychology study program completed a questionnaire about several admission methods. General favorability, ratings on justice dimensions, relationships between general favorability and these dimensions, and differences in perceptions based on gender and on the aim of the admission procedure (selection or matching) were studied. In addition, the relationship between favorability and test performance, and the relationship between favorability and behavioral outcomes were investigated. Applicants rated interviews and trial-studying tests most favorably. Contrary to expectations based on the existing literature, high school grades were perceived least favorably and there was no relationship between applicant perceptions and enrollment decisions. In line with previous research in the employment literature, general favorability was most strongly related to face validity, study-relatedness, applicant differentiation, the chance to show skills, perceived scientific evidence, and perceived wide-spread use. We found no differences in applicant perceptions based on gender and small differences based on the aim of admission procedures. These results extend the applicant perceptions literature to educational admission and the results are useful for administrators when choosing methods to admit students.

In recent years there has been an increasing interest in the use of nontraditional instruments for admission into higher education, such as the use of personality questionnaires, motivation questionnaires, biodata, and trial-studying tests (Niessen, Meijer, & Tendeiro, 2016; Schmitt, 2012; Visser, van der Maas, Engels-Freeke, & Vorst, 2012). Through the administration of these instruments as alternatives for or in addition to traditional entrance exams and high school Grade Point Average (GPA), a broader set of characteristics and skills can be evaluated than using the traditional cognition-based methods (e.g., Lievens & Coetsier, 2002; Schmitt, 2012; Schultz & Zedeck, 2012). Most studies have focused on the effectiveness of these instruments from the perspective of the educational institutions by studying predictive validity and differences between relevant groups. Although such studies are important and show practically and theoretically relevant results, very little attention has been paid to applicant perceptions of different admission methods. Applicant perceptions of selection methods have been mainly studied in the context of personnel selection. However, with the increasing interest in the use of different admission methods in higher education, I/O psychologists, selection officers, and other professionals are

confronted with the question which methods are preferred by candidates in educational admission (e.g., Schmitt, 2012). In the present study, we tried to answer this question by investigating applicant perceptions of different admission methods in higher education, and by investigating relationships of applicant perceptions with test performance and future behavior. In addition, we studied differences in applicant perceptions and admission method preferences for male and female applicants, and differences in admission method preferences depending on the aim of the admission procedure; selection (high-stakes), or matching (low-stakes).

1.1 | Applicant perceptions

Applicant perceptions are "attitudes, affect, or cognitions an individual might have about a selection process" (Ryan & Ployhart, 2000, p. 566), and these perceptions have been widely studied in the context of personnel selection. Different models (Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Gilliland, 1993; Ryan & Ployhart, 2000) and instruments (Bauer, Truxillo, Sanchez, Craig, Ferrara, & Campion, 2001; Sanchez, Truxillo, & Bauer, 2000; Steiner & Gilliland, 1996) have been

developed and consequences of applicant perceptions have been studied. Results showed that applicant perceptions of selection methods are related to test validity, organizational attractiveness, application recommendations to others, job-offer acceptance, litigation likelihood, applicant withdrawal, and purchase intentions (Bauer et al., 2001; Gilliland, 1994; Hausknecht, Day, & Thomas, 2004; Macan, Avedon, Paese, & Smith, 1994; Ryan, Sacco, McFarland, & Kriska, 2000; Smither, Reilly, Millsap, & Pearlman, 1993; Thorsteinson & Ryan, 1997; Truxillo, Steiner, & Gilliland, 2004).

Many of these outcomes are, *mutatis mutandis*, also important for educational institutes. Moreover, higher educational institutes serve important societal purposes and the opportunity to participate in higher education has a large impact on the careers and thus future lives of individuals. Because of the impact of higher education on society and individuals, the perceptions of stakeholders to selection methods are of great importance. An example is the ongoing public debate about the content and importance of the SAT in college admissions and the recent changes made to increase relevance and face validity (e.g., Balf, 2014). Furthermore, it is not self-evident that results based on studies conducted in personnel selection contexts can be generalized to the context of admission to higher education. The outcomes to be predicted in both contexts differ; in personnel selection the main outcome to be predicted is job performance, whereas in educational selection it is academic performance. These different outcomes are predicted by partly different instruments or methods. Some instruments are used in both contexts (e.g., cognitive ability tests, personality questionnaires), but other frequently used admission methods are unique to the context of higher education (high school GPA, lottery). Furthermore, the popularity of different methods may differ across the two contexts (e.g., Ryan, McFarland, Baron, & Page, 1999).

1.2 | Theoretical framework

The dominant perspective on applicant perceptions of selection methods is based on organizational justice theory (Gilliland, 1993). Within organizational justice theory, several *procedural justice dimensions* are proposed that explain the *process favorability* of selection methods. Procedural justice concerns the procedures used to determine the best applicants, opposed to distributive justice, which is focused on the outcomes of the selection procedures (Steiner & Gilliland, 2001). Process favorability (a general preference for a selection method), is determined by perceived fairness and perceived predictive validity (Smither et al., 1993; Steiner & Gilliland, 1996). The seven proposed dimensions of procedural justice are *scientific evidence*, *the right to obtain information*, *applicant differentiation*, *interpersonal warmth*, *face validity*, *wide-spread use*, and *respect of privacy*. These dimensions are usually measured with single items (Steiner & Gilliland, 1996). The organizational justice perspective was supported by findings in many studies (e.g., Schmitt, Oswald, Kim, Gillespie, & Ramsay, 2004; Smither et al., 1993). In the remainder of this article, we shorten the terms *process favorability* and *procedural justice dimensions* to *general favorability* and *justice dimensions* for simplicity.

Sanchez et al. (2000) proposed an alternative perspective on applicant perceptions based on expectancy theory. The three major components of expectancy theory are valence (the desirability of the outcome), instrumentality (the belief that good performance will lead to the desired outcome), and expectancy (the subjective belief that effort will increase the chance of the desired outcome). Sanchez et al. (2000) proposed that these components might partly explain test-taking motivation and procedural justice perceptions.

Another possible determinant of applicant perceptions is the self-serving bias (Chan et al., 1997; Chan, Schmitt, Sacco, & DeShon, 1998; Schmitt et al., 2004). According to this theory, applicants who perform poorly attribute those results to a lack of relevance and fairness of the test. In the studies cited above, small to moderate positive relationships were found between test scores and post-test applicant perceptions, even when controlling for pretest applicant perceptions (Chan et al., 1998).

A more specific characteristic of some methods that has received much attention but has rarely been studied in relation to applicant perceptions is the fakeability or cheatability of selection methods. Many nontraditional methods that are currently receiving attention measure typical behavior (e.g., personality questionnaires, situational judgment tests (SJTs), biodata). These types of tests are susceptible to cheating or faking when used in maximum performance contexts such as selection situations (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006; Viswesvaran & Ones, 1999). Some studies showed that the perceived fakeability of methods was related to applicant perceptions (Gilliland, 1995; Schreurs, Derous, Proost, Notelaers, & de Witte, 2008).

1.3 | Applicant perceptions in personnel selection

Many studies on applicant perceptions have been conducted in the context of personnel selection (e.g. Anderson & Witvliet, 2008; Gilliland, 1994; Smither et al., 1993). Anderson, Salgado, and Hülsheger (2010) conducted a meta-analysis on applicant perceptions using data from many different countries. They found that applicant perceptions were generalizable across specific selection situations and countries. In general, work samples and interviews were the most favorable methods, *résumés*, cognitive ability tests, references, biodata, and personality questionnaires were rated favorably, and honesty tests, personal contacts, and graphology were rated least favorably. Anderson et al. (2010) also found that for the more specific justice dimensions, work samples and interviews were perceived as highly face-valid and were rated favorably on most dimensions. However, work samples were rated slightly lower on *interpersonal warmth*, *scientific evidence*, and *wide-spread use*. Cognitive ability tests were rated highest for *respect of privacy*, and personality tests and biodata were rated moderately on most dimensions.

Relationships between ratings on the justice dimensions and general favorability have been studied to gain insight in the determinants of applicant perceptions. The results were mostly consistent across studies and showed that *face validity*, *applicant differentiation* and *wide-spread use* were strongly related to general favorability, the *right to use* and *scientific evidence* were moderately related to general favorability,

and *interpersonal warmth* and *respect of privacy* showed small relations to general favorability (Bertolino & Steiner, 2007; Ispas, Ilie, Iliescu, Johnson, & Harris, 2010; Moscoso & Salgado, 2004; Nikolaou & Judge, 2007; Steiner & Gilliland, 1996). Another dimension that was strongly related to general favorability but that was not included in Steiner and Gilliland's (1996) framework was *job-relatedness* (Bauer et al., 2001).

In conclusion, high-fidelity methods (methods that are similar to the criterion in content) like work samples, and methods that make applicants feel that they can show their unique skills and abilities, like interviews, are perceived favorably by applicants (e.g. Ployhart, Schneider & Schmitt, 2006).

1.4 | Applicant perceptions in higher education

In the context of higher education, few studies on applicant perceptions of admission methods have been conducted, and the available studies only evaluated specific admission instruments and specific aspects of applicant perceptions. Patterson, Zibarras, Carr, Irish, and Gregory (2011) found that applicants to a post-graduate medical training program rated a clinical problem-solving task as significantly more relevant than a SJT, and a simulated patient task as significantly more relevant than a group exercise and a written exercise. Lievens (2013) found that medical school applicants rated an SJT measuring interpersonal skills as significantly more face valid than cognitive science knowledge tests. These results showed that methods that matched the context of the programs were rated more positively than more general or low-fidelity methods (Kluger & Rothstein, 1993; Ployhart et al., 2006).

In contrast, Schmitt et al. (2004) studied fairness and relevance perceptions of undergraduate students to SAT/ACT scores, and a combined biodata/SJT instrument designed to predict broad college student performance criteria. They found that fairness perceptions for SAT/ACT were higher than for the SJT and biodata instruments, and that fairness ratings were low for the latter two methods. There were no significant differences between the methods for perceived relevance. Schmitt et al. (2004) also studied the effect of direct or indirect self-serving bias and found that perceived performance was positively related to perceptions of relevance, which in turn were positively related to fairness perceptions. Finally, Schmitt (2012) discussed that their "previous collection of reactions measures suggests that students view HSGPA as the most appropriate index of student potential with the use of biodata, SJT, and SAT/ACT less favorably viewed. The latter three indices were perceived to be about equally relevant and fair" (p. 28).

1.5 | Potential variables affecting applicant perceptions

It is well known that in higher education selection performance on some predictors differs across males and females, and that some predictors show differential prediction by gender (Fischer, Schult, & Hell, 2013; Keiser, Sackett, Kuncel, & Brothen, 2016). Males tend to obtain higher scores on cognitive tests than females, and female academic

performance tends to be slightly underpredicted by scores on cognitive tests such as the SAT and ACT (Fisher et al., 2013). Conversely, females tend to score higher on relevant personality constructs such as conscientiousness, procrastination (reversed), and academic skills (Keiser et al., 2016). Therefore, applicant perceptions of admission methods may differ for males and females.

Furthermore, the admission ratio of universities can differ widely. Some admission procedures are aimed at strict *selection* and thus admission of the best candidates, while other procedures are aimed at determining student-program fit (matching), resulting in an *enrollment advice*. Applicant perceptions of admission methods may differ depending on the aim of admission procedures. Some methods may be perceived more favorably when they are used to determine which applicants would be the most successful students (selection), while others may be perceived more favorably when they are used to gain insight in applicants' fit to a program (matching).

1.6 | Aims of the present study

Educational institutes can often choose their own admission methods and criteria to select students, and there is wide variety of possible methods and instruments. Knowledge about perceptions of applicants to higher education about these methods is lacking, and through this study we aimed to fill this gap. Educational institutes can then take this information into account in designing their admission policies. In addition, we investigated if results based on organizational justice theory obtained in an educational context and applied to educational admission methods are comparable to results obtained in personnel selection contexts. We also investigated if applicant perceptions differed depending on gender or on the aim of admission procedures.

After a long tradition of open admission and lottery admission, selective admission was recently implemented in the Netherlands. We studied applicant perceptions of methods that are often used or suggested in the literature, or have recently been implemented in admission to Dutch higher education, based on inspection of websites of higher education institutions (ISO, 2014). These methods were cognitive ability tests, personality questionnaires, motivation questionnaires, biodata, high school grades, subject tests, trial-studying tests, interviews, and lotteries. Table 1 provides a brief description of each method.

First, we studied the general favorability of the admission methods in a selection and a matching sample. We hypothesized that interviews and high-fidelity methods like trial-studying and subject tests would be perceived as most favorable, followed by cognitive ability tests, high school grades, and biodata; lotteries would be perceived as least favorably. Second, we studied ratings on several justice dimensions for each of the methods and their relationships with general favorability to gain insight in determinants of applicant perceptions in higher education. Third, we examined whether applicants perceptions of admission methods differed based on gender, and on the aim of the admission procedure (selection or matching). Fourth, we studied the relationships between general favorability of subject tests and trial-studying tests, and actual test scores obtained with these methods. On the basis of

TABLE 1 Surveyed selection methods and descriptions

Method	Description
Trial-studying ^{a,b}	In trial-studying a part of the study program (mostly the first course) is mimicked. Students complete an exam or assignment very similar to an exam or assignment in the actual program.
Subject tests ^a	Subject tests assess specific skills and abilities on a subject that is very relevant for the discipline of interest.
Personality questionnaires ^c	In personality questionnaires you are asked to respond to statements about yourself to assess your personality traits. An example statement is: <i>I am a hard worker</i> (Strongly disagree—Strongly agree)
Motivation questionnaires ^c	In motivation questionnaires you are asked to respond to statements about yourself to assess your motivation. An example statement is: <i>In my study, my goal is to do better than I did before.</i> (Strongly disagree—Strongly agree)
Cognitive ability tests	Cognitive ability tests are tests that evaluate your intelligence on your reasoning, verbal skills, or mathematical skills.
High school grades	High school grades are used to assess how well you performed in high school.
Biodata	Biodata give an extensive description of all your work experience and education, often including skills, abilities, references and reflections.
Interviews	An interview is a face-to-face interaction in which an admissions officer or employee of the university asks you a variety of questions about your background, skills, and motivation.
Lottery	Some universities base their admission decisions on weighted lotteries. Each applicant is placed in 1 of 5 lottery categories based on their average high school grade. The higher the grade (and the category), the larger the chance of being admitted.

Notes.^aAll participants in the selection sample were evaluated with these methods.^bAll participants in the matching sample were evaluated with this method.^cSixty-one percent of the respondents in the selection sample completed these instruments for research purposes.

the self-serving bias theory we expected that applicants with lower scores would rate the methods as less favorably. Finally, we tried to replicate the relationship between applicant perceptions and behavioral outcomes such as job-offer acceptance found in personnel selection contexts (Hausknecht et al., 2004; Macan et al., 1994). We analyzed the relationship between applicant perceptions of the methods used in an admission procedure and enrollment decisions and we asked the applicants if they took the admission method into account when choosing a university and a study program.

2 | METHOD

2.1 | Participants

2.1.1 | Selection sample

The sample consisted of 220 applicants to an undergraduate psychology program at a Dutch university in 2015. Before participating in the study, the applicants participated in a selection procedure consisting of two trial-studying tests and a subject test in mathematics. The trial-studying tests mimicked future study behavior. For the first trial-studying test applicants were asked to study two chapters of introductory psychology material, and for the second trial-studying test applicants were instructed to view a video lecture. Both tests consisted of multiple-choice questions about the material. The subject test in math consisted of items about high-school algebra and skills related to basic statistics. The selection committee rejected none of the applicants.

However, the students did not know this in advance and perceived the selection tests as high-stakes tests. In addition, 134 of the 220 participants (61%) also voluntarily completed personality and motivation questionnaires for research purposes before participating in the selection procedure. After the selection procedure all applicants were asked to complete an online questionnaire about different selection methods. Participation was voluntary, 34% of all applicants completed the questionnaire. Some participants completed the questionnaire after receiving their scores (22% of the participants). Participants applied to a Dutch-spoken program (34% of the participants, 35% in the applicant pool) or to an English-spoken program. For this latter program mostly international students applied, of which 98% had a European nationality. In the group of participants, 75% was female (70% in the applicant pool). The mean age for the participants was $M = 20$ ($SD = 2.3$), and in the total applicant pool the mean age was $M = 20$ ($SD = 2.2$). Ten percent of the participants decided not to enroll in the program after acceptance to the program (27% in the applicant pool).

2.1.2 | Matching sample

The sample consisted of 133 applicants to the same undergraduate psychology program at a Dutch university in 2016. The faculty had abolished selective admission and implemented a matching procedure instead, that consisted of the same trial-studying tests as in 2015. In addition, the math test was replaced by another trial-studying test about statistics, which covers a significant proportion of the curriculum. The matching procedure was aimed at helping the applicants gain insight into their fit to

TABLE 2 Mean scores, standard deviations, and 95% confidence intervals for general favorability ratings obtained in the selection and the matching sample, and Cohen's *d* for the difference in ratings between the matching sample and the selection sample

Method	Selection			Matching			<i>d</i>
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI	
Interviews	5.29	1.12	[5.15, 5.45] ^a	4.91	1.14	[4.71, 5.10] ^a	-.37*
Trial-studying tests	5.16	1.05	[5.03, 5.31] ^a	4.65	1.12	[4.46, 4.85] ^a	-.48*
Cognitive ability tests	4.72	1.21	[4.56, 4.89] ^b	4.61	1.13	[4.41, 4.80] ^a	-.12
Subject tests	4.70	1.16	[4.53, 4.84] ^b	4.77	1.13	[4.57, 4.96] ^a	.05
Biodata	4.39	1.40	[4.22, 4.59] ^{b,c}	3.79	1.26	[3.58, 4.01] ^b	-.47*
Motivation questionnaires	4.15	1.50	[3.96, 4.36] ^{c,d}	4.15	1.32	[3.92, 4.37] ^b	.01
Personality questionnaires	3.81	1.40	[3.64, 4.01] ^d	3.97	1.26	[3.75, 4.19] ^b	.11
High school GPA	3.28	1.38	[3.11, 3.47] ^e	3.09	1.34	[2.86, 3.32] ^c	-.12
Lottery	3.06	1.29	[2.89, 3.23] ^e				

Notes. Letters in superscript show overlapping confidence intervals. **p* < .05.

the program. The applicants knew that they could not be rejected, but that they would be advised about their enrollment based on the scores on the admission tests. After completing the matching tests all applicants were asked to complete an online questionnaire about different admission methods. Participation was voluntary, 29% of all applicants completed the questionnaire. Participants applied to a Dutch-spoken program (51% of the participants, 47% in the applicant pool) or to an English-spoken program. For this latter program mostly international students applied, of which 86% was from Europe. In the group of participants, 71% was female (70% in the applicant pool). The mean age for the participants was *M* = 20 (*SD* = 3.7), and in the total applicant pool the mean age was *M* = 20 (*SD* = 2.9).

2.2 | Measures

Participants completed an online questionnaire about all admission methods listed in Table 1. For the matching sample, lottery was not included because lottery would not be used for assessing student-program fit. The order of presenting the methods to the respondents was randomly generated for each respondent. Each method was briefly described, sometimes including an example item. Next, 13 items were administered that were mostly based on the questionnaire by Steiner and Gilliland (1996). The first two items (perceived predictive validity and perceived fairness) measured general favorability, and lead to an overall description of the favorability of the methods. In addition, the seven items from this questionnaire measuring justice dimensions were included. We extended the questionnaire with an item about *study-relatedness*, and an item about the *chance to perform* based on Bauer et al. (2001), a question about *effort expectancy* from Sanchez et al. (2000), and a question about the *ease of cheating*. The complete questionnaire can be found in the Appendix Table A1. Each response was provided on a seven-point scale (scored 1–7) with verbal anchors. The respondents completed the questionnaire in Dutch when they applied for the Dutch-spoken program and in English when they applied for

the English-spoken program. In addition, we also asked the participants whether the selection or matching procedure used by a particular university influenced their application for a university and study program (yes, somewhat, or no). Test performance, enrollment in the program, and gender were obtained through the university administration. Informed consent was obtained from all participants to access their test scores and academic records and to match these scores and records with their responses on the questionnaire.

2.3 | Procedure

For both samples, general favorability of each method for each respondent was calculated as the mean score on the two general favorability items. These mean scores were used to calculate the mean favorability and a 95% confidence interval for each selection method. The items measuring interpersonal warmth were reverse scored to ease interpretation. Mean scores and confidence intervals for the justice dimension items were also computed for all admission methods. To study relationships between general favorability and the justice dimensions, we calculated the correlation between scores on the dimension items and the mean general favorability score for each method. To investigate self-serving bias, we computed correlations between the admission test scores and the general favorability ratings of the corresponding method. A logistic regression analysis was conducted with enrollment to the program as the dependent variable and the favorability ratings of trial-studying and subject tests as the independent variables, based on the data obtained in the selection sample. There were 0.4% missing values in the data of the selection sample and no missing data in the matching sample. Since the percentage of missing values was very small and no patterns emerged in the missing data, we made the assumption that the data were missing completely at random and we used pairwise deletion for all analyses. To study if applicant perceptions differed depending on the aim of the admission procedure and the gender of the applicants, a repeated measures

ANOVA was conducted on a dataset containing data from both samples, with the mean general favorability rating as the dependent variable, method as a within-subjects independent variable, and aim and gender as between-subjects independent variables, including an interaction terms between method and aim and method and gender. All analyses were conducted using SPSS version 23.

3 | RESULTS

3.1 | General favorability

First, we assessed if there were differences between participants who completed the questionnaire before or after receiving their admission scores, and between participants who did and who did not complete the personality and motivation questionnaires in the selection sample. We found no differences in favorability ratings between participants who completed the questionnaire before or after receiving their scores on the methods used in the admission procedure, with Cohen's $d = .01$, $t_{(218)} = -.08$, $p = .93$ for trial-studying tests, and Cohen's $d = .20$, $t_{(218)} = 1.24$, $p = .22$ for subject tests. We also found no differences in favorability ratings of personality questionnaires and motivation questionnaires between respondents who completed these instruments and respondents who did not, with Cohen's $d = .07$, $t_{(217)} = .47$, $p = .64$ for personality questionnaires, and Cohen's $d = .14$, $t_{(217)} = .96$, $p = .34$ for motivation questionnaires. Given these results, we combined all cases as a single selection sample.

Table 2 shows descriptive statistics of the general favorability ratings of each method in both samples. In the selection sample, interviews and trial-studying tests received the highest ratings, with nonoverlapping confidence intervals with other methods. Cognitive ability tests, subject tests, biodata, motivation questionnaires, and personality questionnaires were rated less favorably, but the confidence intervals were above or included the neutral mid-point of the scale. High school grades and lotteries were rated least favorably, with nonoverlapping confidence intervals with ratings of the other methods or with the midpoint of the scale. The results in the matching sample were similar but showed a slightly different ordering, with interviews, subject tests, trial-studying tests, and cognitive ability tests rated as most favorable, followed by motivation questionnaires, personality questionnaires, and biodata. High school grades were rated least favorably, with nonoverlapping confidence intervals with other methods. The most salient result in both samples was the low rating of the use of high school grades. Although frequently used and strongly related to academic performance, students did not perceive high school grades as a favorable basis for selection decisions.

3.2 | Justice dimensions

Table 3 shows the scores on all dimensions for each method based on the selection sample. The dimensions *right to use* and *wide-spread use* showed very small differences between the methods. For *invasion of privacy* there were also few differences, and none of the methods were rated as invasive. The dimensions that showed most variation between

methods were *interpersonal warmth*, *applicant differentiation*, *ease of cheating*, *effort expectancy*, and *chance to perform*. None of the methods were rated highly on *study-relatedness*, with the highest mean ratings around the midpoint of the scale. Trial-studying tests scored high on most positive dimensions, but were also perceived as impersonal. Interviews also scored high on most positive dimension and were perceived as personal, as expected. Trial-studying tests and cognitive ability tests scored highest on *scientific evidence*. Lotteries received the lowest scores on all positive dimensions, but also scored low on *ease of cheating*. The most salient results were, again, the unexpected low ratings for high school grades and the mid-range scores for trial-studying on *chance to perform*, *study-relatedness* and *applicant differentiation*, which were lower than expected. Nontraditional measures often used to measure noncognitive skills were rated highest on *ease of cheating*, but were rated favorably for *interpersonal warmth*.

Table 3 also displays the correlations between the dimension ratings and general favorability for all methods. The ratings on *face validity* were most strongly related to general favorability and this relationship was large for all methods. Other strong relationships with general favorability were found for *study-relatedness*, *applicant differentiation*, *chance to perform*, *scientific evidence*, and *wide-spread use*. *Right to use*, *interpersonal warmth*, and *effort expectancy* showed small positive or no relationships with general favorability, and these relationships varied across methods. As expected, *invasion of privacy* showed negative relationships with general favorability, but these relationships were mostly small and not significant. A notable result was the negative correlation between *effort expectancy* and general favorability for personality questionnaires and motivation questionnaires. This may be explained by the possibility of faking on these methods. The dimension *ease of cheating* showed varying relationships with general favorability across methods. Especially motivation questionnaires were rated less favorably when they were rated as easier to fake, as were personality tests and interviews. Previous findings that *face validity* and *job/study-relatedness* were strongly related to general favorability were thus replicated. The same analyses were also conducted for the matching sample and showed very similar results (not tabulated). The most notable differences were seen in the ratings on *study-relatedness*. Subject tests were rated as most study-related in the matching sample, while they were ranked sixth on *study-relatedness* in the selection sample. Cognitive ability tests were rated as most study-related in the selection sample, while they were ranked seventh on *study-relatedness* in the matching sample. Detailed results can be obtained from the first author.

3.3 | Differences in applicant perceptions based on aim and gender

Table 2 shows descriptive statistics for general favorability ratings of the admission methods for selection and matching purposes and Table 4 shows descriptive statistics for males and females in both samples. A repeated measures ANOVA was conducted to investigate if there were differences in favorability ratings depending on the aim of the admission procedure and depending on the gender of the applicants. Mauchly's test showed that sphericity was violated with $\epsilon > .75$, so the

TABLE 3 Mean scores, standard deviations, 95% confidence intervals, and correlations with general favorability for each method on each dimension in the selection sample, with ratings in descending order per dimension

Dimension	Method	M	SD	95% CI	r
<i>Face validity</i>	Overall				.64
	Trial-studying	5.17	1.37	[4.99, 5.35] ^a	.55
	Interviews	5.04	1.37	[4.85, 5.22] ^{a,b}	.62
	Subject tests	4.74	1.34	[4.56, 4.91] ^{b,c}	.65
	Cognitive ability tests	4.64	1.37	[4.46, 4.82] ^{c,d}	.67
	Biodata	4.30	1.54	[4.10, 4.50] ^{d,e}	.72
	Motivation questionnaires	4.21	1.64	[3.99, 4.43] ^{e,f}	.75
	Personality questionnaires	3.86	1.59	[3.65, 4.07] ^f	.60
	High school GPA	3.29	1.63	[3.08, 3.51] ^g	.65
	Lottery	2.52	1.38	[2.34, 2.70] ^h	.50
<i>Applicant differentiation</i>	Overall				.50
	Interviews	5.42	1.30	[5.25, 5.59] ^a	.66
	Biodata	4.90	1.44	[4.71, 5.10] ^b	.54
	Cognitive ability tests	4.77	1.44	[4.58, 4.96] ^b	.58
	Personality questionnaires	4.66	1.62	[4.45, 4.88] ^b	.53
	Motivation questionnaires	4.11	1.60	[3.90, 4.32] ^c	.68
	Subject tests	4.01	1.56	[3.81, 4.22] ^{c,d}	.30
	Trial-studying	3.64	1.57	[3.43, 3.85] ^d	.20
	High school GPA	3.15	1.62	[2.93, 3.36] ^e	.52
	Lottery	2.07	1.31	[1.89, 2.24] ^f	.35
<i>Study-relatedness</i>	Overall				.49
	Cognitive ability tests	3.85	1.42	[3.66, 4.03] ^a	.56
	Interviews	3.73	1.15	[3.73, 4.13] ^a	.44
	Motivation questionnaires	3.51	1.65	[3.29, 3.73] ^{a,b}	.66
	Trial-studying	3.46	1.44	[3.27, 3.65] ^b	.40
	Biodata	3.36	1.48	[3.17, 3.56] ^{b,c}	.51
	Subject tests	3.34	1.43	[3.15, 3.53] ^{b,c}	.46
	Personality questionnaires	3.05	1.49	[2.85, 3.24] ^c	.54
	High school GPA	2.63	1.43	[2.44, 2.82] ^d	.54
	Lottery	2.39	1.37	[2.21, 2.58] ^d	.27
<i>Chance to perform</i>	Overall				.48
	Interviews	4.88	1.51	[4.67, 5.08] ^a	.56
	Biodata	4.70	1.50	[4.50, 4.90] ^a	.50
	Cognitive ability tests	4.63	1.43	[4.44, 4.82] ^a	.59
	Subject tests	4.03	1.49	[3.83, 4.23] ^b	.43
	Personality questionnaires	4.02	1.70	[3.80, 4.25] ^b	.40
	Motivation questionnaires	3.91	1.67	[3.69, 4.13] ^b	.61
	Trial-studying	3.82	1.46	[3.63, 4.01] ^b	.36
	High school GPA	3.20	1.62	[2.98, 3.42] ^c	.51
	Lottery	1.95	1.27	[1.78, 2.11] ^d	.30
<i>Scientific evidence</i>	Overall				.44
	Trial-studying	4.87	1.09	[4.72, 5.01] ^a	.36
	Cognitive ability tests	4.82	1.23	[4.66, 4.99] ^{a,b}	.45
	Subject tests	4.55	1.24	[4.39, 4.71] ^{b,c}	.41
	Interviews	4.24	1.30	[4.06, 4.41] ^c	.34
	Personality questionnaires	3.76	1.37	[3.60, 3.97] ^d	.41
	Biodata	3.75	1.28	[3.58, 3.92] ^{d,e}	.53
	Motivation questionnaires	3.67	1.28	[3.50, 3.83] ^{d,e}	.48
	High school GPA	3.40	1.40	[3.22, 3.59] ^e	.52
	Lottery	2.85	1.38	[2.66, 3.03] ^f	.43
<i>Widely used</i>	Overall				.42
	Trial-studying	4.71	1.30	[4.54, 4.88] ^a	.26
	Subject tests	4.62	1.29	[4.45, 4.79] ^a	.32
	Interviews	4.54	1.31	[4.37, 4.72] ^a	.31
	Cognitive ability tests	4.20	1.23	[4.04, 4.36] ^b	.44
	Motivation questionnaires	3.95	1.36	[3.77, 4.13] ^{b,c}	.58
	Biodata	3.87	1.33	[3.70, 4.05] ^{b,c}	.49
	Personality questionnaires	3.70	1.23	[3.53, 3.86] ^c	.47
	High school GPA	3.60	1.55	[3.40, 3.81] ^{c,d}	.47
	Lottery	3.33	1.42	[3.14, 3.51] ^d	.41

(continues)

TABLE 3 (continued)

Dimension	Method	M	SD	95% CI	<i>r</i>
<i>Right to use</i>	Overall				.23
	Trial-studying	5.33	1.24	[5.16, 5.49] ^a	.15
	Subject tests	5.28	1.30	[5.11, 5.45] ^{a,b}	.09
	Interviews	5.27	1.22	[5.11, 5.43] ^{a,b}	.35
	Motivation questionnaires	4.99	1.20	[4.83, 5.15] ^{b,c}	.20
	Biodata	4.95	1.35	[4.77, 5.13] ^{b,c}	.28
	Cognitive ability tests	4.92	1.28	[4.75, 5.09] ^{c,d}	.39
	High school GPA	4.90	1.34	[4.73, 5.08] ^{c,d}	.29
	Personality questionnaires	4.72	1.45	[4.53, 4.91] ^{c,d}	.25
	Lottery	4.57	1.48	[4.37, 4.76] ^d	.06
<i>Ease of cheating</i>	Overall				−.15
	Motivation questionnaires	5.48	1.58	[5.27, 5.69] ^a	−.48
	Personality questionnaires	5.27	1.87	[5.04, 5.54] ^a	−.25
	Biodata	4.23	1.70	[4.01, 4.45] ^b	−.11
	Interviews	3.18	1.86	[3.56, 4.06] ^b	−.28
	Trial-studying	2.97	1.37	[2.79, 3.15] ^c	−.14
	Subject tests	2.79	1.38	[2.61, 2.98] ^c	.03
	High school GPA	2.67	1.55	[2.46, 2.88] ^c	−.03
	Cognitive ability tests	2.48	1.21	[2.48, 2.80] ^c	−.10
	Lottery	2.05	1.27	[1.89, 2.22] ^d	.03
<i>Effort expectancy</i>	Overall				.14
	Trial-studying	5.82	1.13	[5.67, 5.97] ^a	.31
	Subject tests	5.37	1.26	[5.20, 5.54] ^b	.22
	High school GPA	5.15	1.43	[4.96, 5.34] ^{b,c}	.22
	Biodata	5.14	1.40	[4.95, 5.32] ^{b,c}	.24
	Motivation questionnaires	4.85	1.68	[4.61, 5.06] ^c	−.09
	Interviews	4.79	1.41	[4.61, 4.98] ^c	.06
	Cognitive ability tests	4.22	1.15	[4.02, 4.42] ^d	.32
	Personality questionnaires	3.67	1.94	[3.41, 3.93] ^d	−.17
	Lottery	2.77	1.83	[2.53, 3.02] ^e	.13
<i>Interpersonal warmth</i>	Overall				.12
	Interviews	6.23	0.98	[6.10, 6.36] ^a	.12
	Personality questionnaires	5.72	1.34	[5.54, 5.90] ^b	.16
	Biodata	5.53	1.23	[5.36, 5.69] ^{b,c}	.06
	Motivation questionnaires	5.20	1.41	[5.02, 5.39] ^c	.30
	Cognitive ability tests	4.15	1.50	[3.95, 4.35] ^d	.11
	High school GPA	3.43	1.75	[3.20, 3.67] ^e	.21
	Subject tests	2.83	1.40	[2.64, 3.01] ^f	.01
	Trial-studying	2.70	1.36	[2.52, 2.88] ^{f,g}	.01
	Lottery	2.42	1.56	[2.21, 2.63] ^g	.05
<i>Invasion of privacy</i>	Overall				−.07
	Personality questionnaires	3.62	1.59	[3.41, 3.83] ^a	−.10
	Biodata	3.12	1.46	[2.93, 3.32] ^b	−.02
	Interviews	3.04	1.44	[2.85, 3.23] ^{b,c}	−.18
	Motivation questionnaires	2.97	1.43	[2.78, 3.16] ^{b,c}	.06
	Cognitive ability tests	2.82	1.37	[2.64, 3.00] ^{b,c}	−.07
	High school GPA	2.75	1.27	[2.58, 2.92] ^{c,d}	−.04
	Lottery	2.34	1.35	[2.16, 2.52] ^{d,e}	−.08
	Subject tests	2.16	1.24	[2.00, 2.33] ^e	−.02
	Trial-studying	2.10	1.15	[1.95, 2.25] ^e	−.19

Notes. Mean correlation between dimension ratings and general favorability are printed in bold. Letters in superscript show overlapping confidence intervals.

Huynh-Feldt correction was applied (Field, 2005). There was a small interaction effect between method and aim ($F_{(6,19, 2135.70)} = 4.92$, $p < .01$, $\eta^2_p = .01$) and a small main effect for aim ($F_{(1, 345)} = 7.62$, $p = .01$, $\eta^2_p = .02$), with lower favorability ratings when the aim was matching, compared to selection. The main effect for method was large ($F_{(6,19, 2135.70)} = 81.38$, $p < .01$, $\eta^2_p = .19$). There was also a small inter-

action effect between method and gender ($F_{(6,19, 2135.70)} = 2.74$, $p = .01$, $\eta^2_p = .01$), but no main effect for gender ($F_{(1, 345)} = 1.87$, $p = .18$, $\eta^2_p = .01$). When inspecting Cohen's *ds* shown in Table 2, we can observe that there were almost no differences in favorability ratings between the two aims, except for trial-studying, biodata, and interviews, which were all rated less favorably when the aim was matching,

TABLE 4 Mean scores, standard deviations for general favorability ratings of males and females in both samples and Cohen's *d* for the difference between ratings by male and female applicants

Method	Aim	Males			Females			<i>d</i>
		<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI	
Interviews	Selection	5.30	1.21	[4.97, 5.63]	5.30	1.10	[5.13, 5.47]	.00
	Matching	4.96	1.19	[4.58, 5.35]	4.89	1.12	[4.66, 5.12]	−.06
Trial-studying tests	Selection	5.08	1.23	[4.75, 5.41]	5.19	0.98	[5.04, 5.35]	.11
	Matching	4.40	1.42	[3.94, 4.86]	4.76	.96	[4.56, 4.96]	.32
Cognitive ability tests	Selection	4.82	1.26	[4.48, 5.17]	4.69	1.18	[4.51, 4.88]	−.11
	Matching	4.74	1.14	[4.37, 5.11]	4.55	1.13	[4.32, 4.78]	−.17
Subject tests	Selection	4.75	1.20	[4.42, 5.05]	4.67	1.17	[4.49, 4.85]	−.07
	Matching	4.63	1.31	[4.20, 5.05]	4.82	1.05	[4.61, 5.04]	.17
Biodata	Selection	4.16	1.31	[3.81, 4.52]	4.49	1.42	[4.27, 4.71]	.24
	Matching	3.91	1.40	[3.46, 4.36]	3.74	1.20	[3.50, 3.99]	−.14
Motivation questionnaires	Selection	3.77	1.60	[3.34, 4.21]	4.29	1.44	[4.07, 4.51]	.35*
	Matching	3.72	1.35	[3.28, 4.16]	4.32	1.28	[4.06, 4.59]	.46*
Personality questionnaires	Selection	3.70	1.64	[3.26, 4.15]	3.86	1.32	[3.66, 4.07]	.11
	Matching	3.95	1.32	[3.52, 4.38]	3.98	1.24	[3.72, 4.23]	.02
High school GPA	Selection	3.38	1.55	[2.96, 3.80]	3.26	1.32	[3.06, 3.46]	−.09
	Matching	3.14	1.56	[2.63, 3.65]	3.07	1.24	[2.82, 3.33]	−.05
Lottery	Selection	2.93	1.31	[2.57, 3.28]	3.10	1.28	[2.90, 3.30]	.13

Note. **p* < .05.

with small to moderate effect sizes. Cohen's *ds* displayed in Table 4 showed the same results. The only method that showed a small difference in favorability based on gender was the motivation questionnaire, receiving higher favorability ratings by females than by males.

3.4 | Applicant perceptions and test scores

In the selection sample we found a positive correlation between the favorability scores of the trial-studying tests and the scores on the first trial-studying tests (study a book): $r = .15$ ($p = .02$). For the second trial-studying test (view a lecture) we found $r = .22$ ($p < .01$), and the correlation between favorability of subject tests and the score on the math test was $r = .19$ ($p = .01$). In the matching samples the math test was replaced by a trial-studying test in statistics for the social sciences. The correlations between the general favorability rating of trial-studying tests and test scores were $r = .26$ ($p < .01$) for the first trial-studying test (study a book), $r = .38$ ($p < .01$) for the second trial-studying test (view a lecture), and $r = .12$ ($p = .17$) for the statistics trial-studying test. So, in general, test scores were positively related to general favorability for that same method, but the effect sizes were small.

3.5 | Applicant perceptions and behavioral outcomes

Participants were asked if the selection method influenced their choice of a university and study program. For choosing a univer-

sity, 20% responded that the selection method influenced their choice, 20% responded that it influenced their choice somewhat, and 60% indicated that it was of no influence. With respect to study program choice, 12% answered that the selection method was of influence, 18% answered that it influenced the choice somewhat, and 70% that it was of no influence. In the matching sample, 8% of the respondents indicated that the matching procedure influenced their choice of a university, 14% reported some influence, and 78% said that it was of no influence for choosing a university. For choosing a program, 5% indicated that the matching procedure influenced their choice, 24% report some influence and 71% report no influence.

Based on the data obtained in the selection sample, a logistic regression analysis was conducted to predict enrollment in the program based on the general favorability ratings of trial-studying and subject tests, since these tests were used in the admission procedure. For trial-studying, the mean rating of applicants who did not enroll was $M = 5.4$ ($SD = 0.98$), and for applicants who did enroll the mean rating was $M = 5.1$ ($SD = 1.05$). For subject tests, the mean rating of applicants who did not enroll was $M = 4.7$ ($SD = 0.97$), and for applicants who did enroll the mean rating was $M = 4.7$ ($SD = 1.20$). The logistic regression model did not significantly predict enrollment (model $\chi^2_{(2)} = 1.02$, $p = .60$), with $OR = 0.78$ (95% CI [0.48; 1.28], Wald $\chi^2 = 0.97$, $p = .33$) for general favorability of trial-studying tests and $OR = 1.07$ (95% CI [0.72; 1.60], Wald $\chi^2 = 0.11$, $p = .75$) for subject tests.

4 | DISCUSSION

The aim of this study was to investigate applicant perceptions of admission methods used in higher education. We found some surprising results; the low favorability of using high school grades for matching or selection purposes was most surprising. High school grades are widely used in many countries and are a highly valid predictor of academic performance in higher education (e.g., Richardson, Abraham, & Bond, 2012). The low favorability of high school grades was contrary to the results found in the personnel selection literature that actual predictive validity was related to general favorability (Anderson et al., 2010), and contrary to Schmitt's (2012) report that high school GPA was viewed most favorably by students and other stakeholders. A possible explanation for our results supported by organizational justice theory (Gilliland, 1993) and expectancy theory (Sanchez et al., 2000) is that high school grades are already obtained and cannot be altered, which may evoke feelings of "not being in control" of the admission process. High school grades were rated low on *chance to perform*, *applicant differentiation*, and *face validity*, which were strongly related to general favorability. The same rationale may apply to the low favorability ratings of lotteries, which were rated least favorably on general favorability and the majority of the justice dimensions.

We also found that the nontraditional methods used to measure noncognitive characteristics (personality and motivation questionnaires and biodata) were not rated very favorably, and significantly less favorably than interviews and trial-studying. These methods were perceived as easy to cheat and the perceived ease of cheating was negatively related to the general favorability of these methods. *Effort expectancy* showed a negative correlation with general favorability for motivation questionnaires and personality questionnaires, while it was positively related to general favorability for all other admission methods. This negative correlation may also be related to the possibility of faking on these methods, where "investing effort" on these methods may have been interpreted as faking by the applicants.

Because of the consistent findings of differences between males and females in scores on cognitive tests and some personality trait measures, we hypothesized that applicant reactions to admission methods may also differ between male and female applicants. We found a significant interaction effect between method and gender, but the effect size was small. In addition, we expected that the aim of the admission procedure (selection or matching) could influence applicant perceptions as well. Our results showed small significant effects for aim and for the interaction between method and aim. Applicants tended to rate methods less favorably when the aim was matching, but these effects were small. A notable finding was that the two most favorably rated methods, interviews and trial-studying, showed relatively large differences in favorability for the selection and matching samples. An explanation for this finding could be that the results of matching procedures are not binding, but that trial-studying tasks and the interviews would require preparation and effort. When applicants have to put effort into a task that does not really have consequences, the result may be a lower appreciation of such a task than when the results would have important consequences.

With respect to the relationships between applicant perceptions and behavioral outcomes there were significant but small correlations between test performance and favorability. In contrast to findings obtained in employment settings (Hausknecht et al., 2004), we found no relationship between applicant perceptions and enrollment decisions. However, these applicants went through an admission procedure consisting of trial-studying tests and a subject test, which were rated favorably. These results might have been different when other, less favorably rated methods were used. Although the majority of applicants in both samples indicated that the admission methods did not influence their choice of a program or a university, between 20 and 40 percent of the applicants indicated that the admission methods influenced their choice at least to some extent. These numbers could be of practical significance to higher education institutions.

4.1 | Limitations

One limitation of this study was that we used two cohorts of applicants to a psychology program at a Dutch university, and that not all applicants participated in the study. However, the participants seemed to be representative for the entire applicant pools, with enrollment rate as an exception. The percentage of participants that chose to enroll in the program was larger than the percentage in the applicant pool. Second, the respondents did not have experience with all admission methods in the questionnaire. In the selection sample, all respondents took trial-studying tests and a subject test, and some respondents also completed a personality and motivation questionnaire for research purposes, but respondents may have differed in the amount of experience they had with other methods. This may have resulted in differences in perceptions between respondents. We did not, however, find differences in perceptions between respondents who did and those who did not complete the personality and motivation questionnaires. In the matching sample, applicants only took three trial-studying tests.

Another possible limitation may be that we used Steiner and Gilliland's (1996) questionnaire that consists of single-item measures for the justice dimensions. Whereas this may lead to reduced validity when measuring broad constructs, single-item measures are suitable for narrow and specific constructs, such as the justice dimensions (e.g. Gardner, Cummings, Dunham, & Pierce, 1998). Also, Jordan and Turner (2008) found that single-items functioned well in measuring organizational justice.

4.2 | Theoretical implications

This study showed that organizational justice theory can be applied to applicant perceptions in an educational context, and this was the first study that applied this theory to a wide variety of admission methods in an educational context. To some extent, we found results similar to the results in personnel selection (e.g., Anderson et al., 2010), with the highest ratings for interviews and trial-studying (a method similar to work sample tests in personnel selection). The favorability of admission methods was most strongly related to their face validity, study-relatedness, differentiation of applicants, the chance to show ones

skills, perceived scientific evidence, and perceived widespread use. In line with previous findings, we found that high-fidelity methods such as trial-studying were rated more favorably than low-fidelity methods such as personality questionnaires. An exception was cognitive ability tests, which is not a very high-fidelity method, but was rated favorably. An explanation for the high favorability of high-fidelity methods is their high face validity and criterion-relatedness (Ployhart et al., 2006). However, trial-studying was not rated highly on study-relatedness in this study.

While organizational justice theory could provide meaningful insight into the favorability of admission methods, the justice dimensions *right to use*, *interpersonal warmth* and *invasion of privacy* that were part of the original applicant perceptions scale by Steiner and Gilliland (1996) showed very little variation in ratings across methods or small correlations with general favorability. These findings are in line with studies conducted in personnel selection contexts (Bertolino & Steiner, 2007; Ispas et al., 2010; Moscoso & Salgado, 2004; Nikolaou & Judge, 2007; Steiner & Gilliland, 1996), although *right to use* was more strongly related to general favorability in those studies. In addition, the dimensions *study-relatedness* and *chance to perform* that we included in the questionnaire used in this study, but obtained from a different instrument developed to measure procedural justice (Bauer et al., 2001) showed strong relationships with general favorability. The effort expectancy dimensions obtained from Sanchez et al., (2000) did not show such a relation. Furthermore, some dimensions that were not included in the original framework may be specifically relevant for some methods. *Ease of cheating* was not related to general favorability for most methods, but it was for self-report instruments. These results indicate the need for reconsidering the procedural justice dimensions that determine the general favorability of admission methods used in education, and perhaps in personnel selection as well.

4.3 | Practical Implications

Applicant perceptions may be taken into account when choosing methods to admit students. However, they should be carefully weighted with the predictive validity of the individual selection methods. The high favorability of interviews, for example, is not in accordance with the often-found low validity and reliability of interviews, especially when they are unstructured (Schmidt & Hunter, 1998). Also, the low favorability of using high school grades does not correspond with their high predictive validity (e.g., Richardson et al., 2012). When methods show similar predictive validity, the method associated with more positive applicant perceptions may be preferred. For example, Niessen et al. (2016) reported high and similar predictive validities for a trial-studying test and high school grades for first-year academic performance. Considering the high favorability of trial-studying tests and the negative applicant perceptions toward using high school grades as an admission criterion, using a trial-studying test may be a viable alternative to high school grades. Alternatively, interventions could be implemented to explain the use of unpopular admission methods so as to influence applicant perceptions (e.g. Truxillo et al., 2009). Furthermore,

more studies that examine the behavioral consequences of positive and negative applicant perceptions are needed.

REFERENCES

- Anderson, N., Salgado, J. F., & Hülsheger, U. R. (2010). Applicant reactions in selection: Comprehensive meta-analysis into reaction generalization versus situational specificity. *International Journal of Selection and Assessment*, 18, 291–304. doi:10.1111/j.1468-2389.2010.00512.x
- Anderson, N., & Witvliet, C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal, and Singapore. *International Journal of Selection and Assessment*, 16, 1–13. doi: 10.1111/j.1468-2389.2008.00404.x
- Balf, T. (2014, March 6). The story behind the SAT overhaul. *The New York Times*, Retrieved from <http://nyti.ms/1cCH2Dz>
- Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, 54, 388–420. doi:10.1111/j.1744-6570.2001.tb00097.x
- Bertolino, M., & Steiner, D. D. (2007). Fairness reactions to selection methods: An Italian study. *International Journal of Selection and Assessment*, 15, 197–205. doi:10.1111/j.1468-2389.2007.00381.x
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14, 317–335. doi:10.1111/j.1468-2389.2006.00354.x
- Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology*, 82, 300–310. doi:10.1037/0021-9010.82.2.300
- Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology*, 83, 471–485. doi:10.1037/0021-9010.83.3.471
- Field, A. P. (2005). *Discovering statistics using SPSS*. London, UK: Sage Publications.
- Fischer, F. T., Schult, J., & Hell, B. (2013). Sex-specific differential prediction of college admission tests: A meta-analysis. *Journal of Educational Psychology*, 105, 478–488. doi:10.1037/a0031956
- Gardner, D. G., Cummings, L. L., Dunham, R. B., & Pierce, J. L. (1998). Single-item versus multiple-item measurement scales: An empirical comparison. *Educational and Psychological Measurement*, 58, 898–915. doi:10.1177/0013164498058006003
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18, 694–734. doi:10.5465/AMR.1993.9402210155
- Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology*, 79, 691–701. doi:10.1037/0021-9010.79.5.691
- Gilliland, S. W. (1995). Fairness from the applicant's perspective: Reactions to employee selection procedures. *International Journal of Selection and Assessment*, 3, 11–19. doi:10.1111/j.1468-2389.1995.tb00002.x
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57, 639–683. doi:10.1111/j.1744-6570.2004.00003.x

- ISO. (2014). *Meer transparantie bij decentrale selectie: Het belang en een framework* [More transparency in admissions: The importance of a framework]. Retrieved from <http://www.iso.nl/website/wp-content/uploads/2014/12/1415-meer-transparantie-bij-decentrale-selectie3.pdf>
- Ispas, D., Ilie, A., Iliescu, D., Johnson, R. E., & Harris, M. M. (2010). Fairness reactions to selection methods: A Romanian study. *International Journal of Selection and Assessment*, 18, 102–110. doi:10.1111/j.1468-2389.2010.00492.x
- Jordan, J. S., & Turner, B. A. (2008). The feasibility of single-item measures for organizational justice. *Measurement in Physical Education and Exercise Science*, 12, 237–257. doi:10.1080/10913670802349790
- Keiser, H. N., Sackett, P. R., Kuncel, N. R., & Brothen, T. (2016). Why women perform better in college than admission scores would predict: Exploring the roles of conscientiousness and course-taking patterns. *Journal of Applied Psychology*, 101, 569–581. doi:10.1037/apl0000069
- Kluger, A. N., & Rothstein, H. R. (1993). The influence of selection test type on applicant reactions to employment testing. *Journal of Business and Psychology*, 8, 3–25. doi:10.1007/BF02230391
- Lievens, F. (2013). Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. *Medical Education*, 47, 182–189. doi:10.1111/medu.12089
- Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment*, 10, 245–257. doi:10.1111/1468-2389.00215
- Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology*, 47, 715–738. doi:10.1111/j.1744-6570.1994.tb01573.x
- Moscato, S., & Salgado, J. F. (2004). Fairness reactions to personnel selection techniques in Spain and Portugal. *International Journal of Selection and Assessment*, 12, 187–196. doi:10.1111/j.0965-075X.2004.00273.x
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2016). Predicting performance in higher education using proximal predictors. *PLoS One*, 11(4), 1–14. doi:10.1371/journal.pone.0153663
- Nikolaou, I., & Judge, T. A. (2007). Fairness reactions to personnel selection techniques in Greece: The role of core self-evaluations. *International Journal of Selection and Assessment*, 15, 206–219. doi:10.1111/j.1468-2389.2007.00382.x
- Patterson, F., Zibarras, L., Carr, V., Irish, B., & Gregory, S. (2011). Evaluating candidate reactions to selection practices using organisational justice theory. *Medical Education*, 45, 289–297. doi:10.1111/j.1365-2923.2010.03808.x
- Ployhart, R. E., Schneider, B., & Schmitt, N. (2006). *Staffing organizations: Contemporary practice and theory*. Mahway, NJ: Lawrence Erlbaum Associates.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138, 353–387. doi:10.1037/a0026838
- Ryan, A. M., McFarland, L. A., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52, 359–391. doi:10.1111/j.1744-6570.1999.tb00165.x
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management*, 26, 565–606. doi:10.1177/014920630002600308
- Ryan, A. M., Sacco, J. M., McFarland, L. A., & Kriska, S. D. (2000). Applicant self-selection: Correlates of withdrawal from a multiple hurdle process. *Journal of Applied Psychology*, 85, 163–179. doi:10.1037/0021-9010.85.2.163
- Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology*, 85, 739–750. doi:10.1037/0021-9010.85.5.739
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274. doi:10.1037/0033-2909.124.2.262
- Schmitt, N. (2012). Development of rationale and measures of noncognitive college student potential. *Educational Psychologist*, 47, 18–29. doi:10.1080/00461520.2011.610680
- Schmitt, N., Oswald, F. L., Kim, B. H., Gillespie, M. A., & Ramsay, L. J. (2004). The impact of justice and self-serving bias explanations of the perceived fairness of different types of selection tests. *International Journal of Selection and Assessment*, 12, 160–171. doi:10.1111/j.0965-075X.2004.00271.x
- Schreurs, B., Derous, E., Proost, K., Notelaers, G., & de Witte, K. (2008). Applicant selection expectations: Validating a multidimensional measure in the military. *International Journal of Selection and Assessment*, 16, 170–176. doi:10.1111/j.1468-2389.2008.00421.x
- Schultz, M. M., & Zedeck, S. (2012). Admission to Law school: New measures. *Educational Psychologist*, 47, 51–65. doi:10.1080/00461520.2011.610679
- Smither, J. W., Reilly, R. R., Millsap, R. E., & Pearlman, K. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49–76. doi:10.1111/j.1744-6570.1993.tb00867.x
- Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology*, 81, 134–141. doi:10.1037/0021-9010.81.2.134
- Steiner, D. D., & Gilliland, S. W. (2001). Procedural justice in personnel selection: International and crosscultural perspectives. *International Journal of Selection and Assessment*, 9, 124–137. doi:10.1111/1468-2389.00169
- Thorsteinson, T. J., & Ryan, A. M. (1997). The effect of selection ratio on the perceptions of the fairness of a selection test battery. *International Journal of Selection and Assessment*, 5, 159–168. doi:10.1111/1468-2389.00056
- Truxillo, D. M., Bodner, T. E., Bertolino, M., Bauer, T. N., & Yonce, C. A. (2009). Effects of explanations on applicant reactions: A meta-analytic review. *International Journal of Selection and Assessment*, 17, 346–361. doi:10.1111/j.1468-2389.2009.00478.x
- Truxillo, D. M., Steiner, D. D., & Gilliland, S. W. (2004). The importance of organizational justice in personnel selection: Defining when selection fairness really matters. *International Journal of Selection and Assessment*, 12, 39–53. doi:10.1111/j.0965-075X.2004.00262.x
- Visser, K., van der Maas, H., Engels-Freeke, M., & Vorst, H. (2012). Het effect op studiesucces van decentrale selectie middels proefstuderen aan de poort [The effect on study success of student selection through trial-studying]. *Tijdschrift Voor Hoger Onderwijs*, 30, 161–173.
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59, 197–210. doi:10.1177/00131649921969802

How to cite this article: Niessen ASM, Meijer RR, Tendeiro JN. Applying organizational justice theory to admission into higher education: Admission from a student perspective. *Int J Select Assess*. 2017;25:72–84. <https://doi.org/10.1111/ijsa.12161>

APPENDIX

TABLE A1 Applicant perceptions questionnaire

Item	General (process) favorability	Source
1. How would you rate the effectiveness of a (method) for identifying qualified people for studying psychology?	Perceived predictive validity	Steiner and Gilliland (1996)
2. If you would <i>not get accepted/receive a negative enrollment advice</i> based on a (method), what would you think of the fairness of this procedure?*	Perceived fairness	Steiner and Gilliland (1996)
(Procedural) justice dimensions		
3. Using a (method) is based on solid scientific research.	Scientific evidence	Steiner and Gilliland (1996)
4. A (method) is a logical test for identifying qualified candidates for studying psychology.	Face validity	Steiner and Gilliland (1996)
5. A (method) will detect an individual's important qualities, differentiating them from others.	Applicant differentiation	Steiner and Gilliland (1996)
6. A (method) is impersonal.	Interpersonal warmth	Steiner and Gilliland (1996)
7. The university has the right to obtain information from applicants by using a (method).	Right to use	Steiner and Gilliland (1996)
8. A (method) invades personal privacy.	Invasion of privacy	Steiner and Gilliland (1996)
9. A (method) is appropriate because methods like this are widely used.	Wide-spread use	Steiner and Gilliland (1996)
10. A person who scores well on a (method) will be a good psychology student	Study-relatedness	Bauer et al. (2001)
11. I could really show my skills and abilities through a (method).	Chance to perform	Bauer et al. (2001)
12. You can get a good score on a (method) if you putt some effort into it.	Effort expectancy	Sanchez et al. (2000)
13. It is easy to cheat or fake on a (method).	Ease of cheating	Self-constructed

Notes. *Phrasing of the question differed for the selection sample and matching sample.